REMARKS

Claims 1-12 are pending and presented for examination in the subject application. Applicants have hereinabove amended independent claims 1 and 7 to place them in better form for consideration, without narrowing the scope of the claimed invention.

Applicants maintain that no new matter and no new issues are presented by this amendment. Accordingly, applicants respectfully request that this Amendment be entered.

Rejection under 35 U.S.C. § 112, second paragraph

In Section 2 of the June 18, 2001 Office Action, claims 1-12 were rejected under 35 U.S.C. §112, second paragraph, as allegedly being indefinite for purportedly failing to particularly point out and distinctly claim the subject matter which applicants regard as the invention.

The June 18, 2001 Office Action stated that claims 1 and 7 were rejected for reciting the phrase "corresponding homologous sequence information", as it purportedly is not clear which sequences would be determined to have "corresponding homology" and the sequences that do not, as would be reasonable for establishing the families. The Office Action also stated that Applicants can overcome this rejection by deleting the term "corresponding".

In response, without conceding the correctness of the rejection, applicants have hereinabove amended claims 1 and 7 in the manner suggested in the Office Action. Applicants maintain that the claim amendments do not change the scope of the claimed invention.

Accordingly, applicants respectfully request reconsideration and withdrawal of the rejection of the claims under 35 U.S.C. § 112, second paragraph.

Rejection Under 35 U.S.C. § 103(a)

In Section 5 of the June 18, 2001 Office Action, claims 1, 3, 5-7, 9, 11 and 12 were rejected under 35 U.S.C. § 103(a) as purportedly being unpatentable over Bachar et al., "A computer vision based technique for 3-D sequence-independent structural comparison of proteins," Protein Engineering, vol. 6, pp. 279-288, 1993 (hereinafter "Bachar paper"), in view of Hendrickson et al., "Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure," The EMBO Journal vol. 9, pp. 1665-1672, 1990 (hereinafter "Hendrickson paper"), in view of Everett et al., "Pendred syndrome is caused by mutations in a putative sulphate transporter gene (PDS)," Nature Genetics, vol. 17, pp. 411-422, 1997 (hereinafter "Everett paper"), and in view of Andrea et al., "Applications of Neural Networks in Quantitative Structure-Activity Relationships of Dihydrofolate Reductase Inhibitors", J. Med. Chem. vol. 34, pp. 2824-2836, 1991 (hereinafter "Andrea paper").

The June 18, 2001 Office Action stated that the Bachar paper teaches a method/system for protein classification, wherein an experimentally-derived, three-dimensional structure of a target protein can be classified by assignment to a cluster set of structurally similar, three-dimensional representation of proteins in an organized database. The Office Action further stated that the design and organization of the database, according to the Bachar paper, consists of three major steps: 1) finding relatively small subset of the structures that form an initial match; 2) finding clusters of initial matches that

represent similar transformations; and 3) extending the clusters to contain additional matching pair residues. The Office Action also stated that these steps are further comprised of sub-steps detailed in the Bachar paper.

The June 18, 2001 Office Action stated that the Bachar paper teaches both a well-known sequence dependent approach in the art in an initial classification step, as well as a sequence independent approach for initial classification steps of target proteins to those of a database. The Office Action further stated that these homology comparison approaches are demonstrated in the Everett paper, wherein the classification of an unknown protein is first carried out through a sequence dependent approach. The Office Action also stated that the Everett paper discloses a method/system for characterizing/clustering proteins into families based on their sequence and structure, such as using their linear sequence and by characterizing transmembrane regions using PHDhtm. The Office Action stated that the Everett paper then assigns a function to an unknown protein based on the similarity comparison of the target protein to the proteins which have been clustered into families which also have functional information. The Office Action further stated that, for example, a function of sulfate transport is assigned to pendrin based on the family clustering model which classified pendrin in the family of other sulfate transporters, and the observed physiological effects that are present which correlate with sulfate transport deficiency in those diagnosed with Pendred syndrome. The Office Action also stated that the Everett paper discloses that these bioinformatics tools are advantageous in that they reduce experimental efforts of trial and error, wherein researchers would otherwise be uncertain of the target protein's function. The Office Action alleged that, combined with the disclosure of the Bachar paper, these limitations of Applicants'

claimed invention are disclosed and obvious, as both the Bachar paper and the Everett paper allegedly recognizes the usefulness of sequence dependent alignment.

The June 18, 2001 Office Action stated that the focus of the Bachar paper is in proteomics wherein protein classification is based on the x-ray crystallographic data, and not in the instrumental improvements of x-ray crystallography. The Office Action alleged that one of ordinary skill in the art would recognize that improved analytical methods for obtaining data with higher precision and/or resolution would be beneficial to the bioinformatics tool of the Bachar paper.

The June 18, 2001 Office Action stated that the Hendrickson paper teaches a system and process for incorporating selenomethionine (as a replacement for methionine) into recombinant proteins produced from plasmids in *E. coli.*, which are crystallized and analyzed by multiwavelength anomalous diffraction (MAD) as a means for producing a three-dimensional representation of a target protein. The Office Action also stated that their method provides the advantages over conventional x-ray techniques for elucidating three-dimensional protein structures, in that MAD utilizes the scattering effects of resonance between x-rays and bound atomic orbitals, it is perfectly isomorphic, allows for data sampling from a single crystal, and the analysis is algebraically exact.

The June 18, 2001 Office Action alleged that one of ordinary skill in the art would have recognized the advantages of the classification system as taught by the Andrea paper with the proteomics tool of the Bachar paper. The Office Action stated that neural networks, as disclosed by the Andrea paper, are dynamic classification systems useful in modeling the non-linear

function of several variables, wherein classification of a changing population of subjects within a database can be carried out. The Office Action alleged that one would recognize that the refined model and structure of the experimentally examined protein would provide an improved and refined overall model, and the physicochemical features of the refined target protein whose structure has been determined could be applied to the other protein family members which were originally classified based on sequence information. The Office Action also alleged that one of ordinary skill in the art would also be motivated to implement the QSARs methodology as disclosed by the Andrea paper, as a means to determine which chemical structures (or family/class of chemical structures) associate with a given protein structure (or family thereof).

Applicants maintain that the Bachar paper, the Hendrickson paper, the Everett paper and the Andrea paper do not render unpatentable the claimed invention set forth in claims 1, 3, 5-7, 9, 11 and 12. The claimed invention is patentable over the Bachar paper, the Hendrickson paper, the Everett paper and the Andrea paper for at least the following reasons.

The subject application claims a novel and unobvious method and system for determining experimentally a plurality of three-dimensional protein structures. Sequence information for a plurality of proteins and structural information and functional information for selected proteins are systematically organized into a database. The sequence information, structural information and functional information stored in the database are used with a bioinformatics tool to cluster the plurality of proteins into families. In each such family, the members have homologous sequences. For each family, a plurality of members of the family are selected as target proteins (for structure

determination). The target proteins are synthesized. Synthesized products are screened, processed and crystallized into specimen crystals. The specimen crystals are tested for predetermined diffraction characteristics to determine suitable specimen crystals. High-throughput X-ray crystallography is performed on the suitable specimen crystals. Diffraction data obtained from the X-ray crystallography is analyzed and used to build and refine an atomic model of the corresponding target protein. The refined model of the target protein is analyzed using sequence information corresponding to other family members and structural information corresponding to other proteins, which are stored in the database, to determine functional motifs and surface characteristics to define active sites and macromolecular contact sites. The subject method and system may be used to develop a comprehensive structural genomics database.

Applicants maintain that the claims of the subject application would not have been obvious to one of ordinary skill in the art at the time of the invention in view of the Bachar paper, considered alone or along with the other cited reference, because the Bachar paper teaches away from the claims.

In order to cover the genomics space, the claims of the subject application provides for systematically organizing the already available information into a database, including in particular the wealth of sequence information for proteins of unknown structure as well as for those of known structure, and using the information with a bioinformatics tool to cluster the proteins into families, in which members of a family have homologous sequences. Using the information in the systematically organized database, a plurality of target proteins may be identified and synthesized for experimental structure determination via high-throughput X-ray crystallography. Diffraction data obtained

through the X-ray crystallography are used to build and refine three-dimensional (3-D) atomic models of the target proteins. The Bachar paper simply does not disclose or suggest, and indeed teaches away from, such a novel and unobvious claimed invention.

The Bachar paper, which is the primary reference cited in the June 18, 2001 Office Action, relates to a method of comparing **known protein structures** in a **sequence-independent** manner. In the method described in the Bachar paper, regions of structural similarity between the protein structures are determined using three-dimensional coordinate data of the known structures to be compared.

The Bachar paper emphasizes that it teaches a **sequence-independent**, "pure" 3-D approach. For example, the Bachar paper states in the section entitled "Discussion" on page 286, in relevant part as follows:

> "... To date, all known methods search for geometric similarities between two proteins where a strict constraint has been placed in the search: sequential order conservation. ...
>
> ... applying such a constraint may be inadequate if the exact evolutionary relationship between the structures is unknown or when possible genetic mutations could have occurred (e.g. an interchange of segments in the sequence). **Structural comparison using such a constraint introduces a sequence-order bias into the results** as it assumes that the structures are evolutionarily related. In addition, sequence-independent structural comparison can help find common three-dimensional folding units (either with or without functional relationship). Thus, **it is important to be able to compare proteins in an unbiased, sequence-independent way**, especially if one is dealing with the question of convergence to a similar structure or divergence from a common ancestor. ..."

The Bachar paper further states in the section entitled "Conclusions" on page 286, in relevant part as follows:

> "... Above all, previous methods are at least partially constrained to linear matches, whereas the 'pure' 3-D approach of our method provides a way to obtain **sequence independent matches** not constrained by the 'progression rule' followed by current alignment techniques. ...
>
> Comparison of protein structures **in a sequence-independent method** provides a way of comparing distantly related and globally dissimilar proteins **without the bias introduced by previous methods of linear** protein structural alignment. Such linear global alignments are sometimes incomplete and inaccurate, whereas our method is capable of discovering partial structural similarities. ...
>
> ... Allowing superpositioning of structural units **independent of their sequential order** enables searches and detection of substructural motifs in the interior of protein molecules."

In view of, for example, the above-quoted portions of the Bachar Paper (see also the title of the Bachar paper: "A computer vision based technique for 3-D sequence-independent structural comparison of proteins"), the Bachar paper clearly teaches the benefits of a strictly three-dimensional approach to protein comparison and teaches away from incorporating sequence information. Thus, the Bachar Paper teaches away from the claimed system/method, wherein, for example, sequence information for a plurality of proteins is used, along with structural information and functional information for selected proteins, to cluster the plurality of proteins into families, in which, for each family, members of the family have homologous sequences.

Applicants further maintain that the claims of this application would not have been obvious to one of ordinary skill in the art at the time of the invention in view of the Bachar paper, considered alone or together with the other cited references, because the exclusive focus of the Bachar paper is on comparisons among those proteins for which three-dimensional structures are already known. This specialized emphasis provides neither the motivation nor methodological basis for implementing pan-genomic

structure determination as provided by the claimed invention set forth in claims 1-12.

The teachings of the Bachar paper relate exclusively to comparisons among proteins of known three-dimensional structure, and they have no realm of application to the structures associated with sequences for which three-dimensional structures have not yet been determined. It is the sequence families associated with unknown structures that must be identified in the process for pan-genomic structure determination, which is the subject of this patent application. Since the methods of the Bachar paper require the very structural information that is being sought in pan-genomic structure determination, they cannot have a bearing on systematically organizing and clustering into families those sequences for which there are no known three-dimensional structures. Thus, the claimed invention in the subject application would not have been obvious from the teachings of the Bachar paper even when considered in combination with any methodological disclosures in the other cited references.

Since the Bachar Paper teaches away from using sequence information to compare and cluster proteins and relates only to that subset of all proteins for which three-dimensional structures are already known, one skilled in the art would not have had motivation at the time of the present invention to combine or adapt the teachings of the Bachar paper with, for example, teachings of use of sequence data in the Everett paper, without impermissibly using the claimed invention as a roadmap.

While teaching away from clustering proteins using the wealth of sequence information available, the Bachar paper does not disclose or suggest a basis for determining three-dimensional

protein structures for which structural information is not already available.

Therefore, it is submitted that the Bachar paper, considered alone or along with other references, would not have rendered the claimed invention set forth in claims 1 and 7 obvious to one of ordinary skill in the art without improper hindsight.

The additional references cited in the Office Action, the Hendrickson paper, the Everett paper and the Andrea paper, do not supply the motivation and the information that the Bachar paper lacks, and therefore do not, alone or in combination, render obvious claim 1 and claim 7.

The Hendrickson paper relates to structural analysis of a select protein. A recombinant selenomethionyl protein (thioredoxin) was expressed in *E.coli*. Selenomethionyl thioredoxins were crystallized and characterized, and then analyzed through X-ray crystallography using a multiwavelength anomalous diffraction (MAD) phasing technique.

The Hendrickson paper does not describe or suggest, however, the benefits of, for example, (a) systematically organizing in a comprehensive database of sequence information for a plurality of proteins and structural information and functional information for selected proteins, (b) using a bioinformatics tool and the sequence information, structural information and functional information in the database to cluster the plurality of proteins into families, in which, for each family, members of the family have homologous sequences, and (c) for each family, selecting a plurality of family members as target proteins for structure determination, as provided by the claimed invention set forth in claims 1 and 7.

The Everett paper relates to a study to identify a gene (PDS) which when mutated is responsible for Pendred syndrome, a recessively inherited disorder with the hallmark features of deafness and thyroid goitre. It was predicted that PDS encodes a 780 amino acid protein (so-called "pendrin").

The Everett paper discloses use of BLAST and PSI-BLAST searches of public sequence databases with the sequence of a protein of interest, pendrin, as the query sequence. The method disclosed in the Everett paper, unlike that of the Bachar paper, is indeed a sequence-dependent approach. It is not, however, a method for characterizing and clustering proteins into families, as contended in the June 18, 2001 Final Office Action; rather, it is a method for finding specific proteins (thirteen in the Everett paper) that are similar in sequence to one query sequence (that of pendrin in the Everett paper). In some circumstances, such as disclosed in the Everett paper, certain structural characteristics such as transmembrane segments may be inferred from features of the sequences, but this is not general. The approach disclosed in the Everett paper does not address the global problem of clustering the universe of known sequences into a comprehensive set of sequence families. Applicants maintain that it would not be obvious to someone of ordinary skill in the art, in view of the disclosure in the Everett paper of such tools as BLAST, PSI-BLAST and PHDhtm, to make such a global clustering into families. Moreover, for reasons described above, neither would such a person be inclined to adopt such a step given the teachings in the Bachar paper, considered in combination with the teachings of the Everett paper.

While the Everett paper discloses a predicted function for a select protein (i.e. pendrin) using sequence information from homologous proteins, the Everett paper, like the Hendrickson

paper, does not disclose or suggest, however, the benefits of, for example, (a) systematically organizing in a comprehensive database of sequence information for a plurality of proteins and structural information and functional information for selected proteins, (b) using the sequence information, structural information and functional information in the database to cluster the plurality of proteins into families, in which, for each family, members of the family have homologous sequences, and (c) for each family, selecting a plurality of family members as target proteins for three-dimensional structure determination, as provided by the claimed invention set forth in claims 1 and 7.

The Andrea paper relates to using a neural network for predicting biological activity of chemical compounds. More specifically, the Andrea paper discusses use of a neural network technique to predict quantitative and qualitative relationships between physicochemical properties and biological activity of a select group of molecules (i.e. dihydrofolate reductase inhibitors).

The Andrea paper, like the Hendrickson paper and the Everett paper, does not describe or suggest, however, the benefits of, for example, (a) systematically organizing in a comprehensive database of sequence information for a plurality of proteins and structural information and functional information for selected proteins, (b) using the sequence information, structural information and functional information in the database to cluster the plurality of proteins into families, in which, for each family, members of the family have homologous sequences, and (c) for each family, selecting a plurality of family members as target proteins for structure determination, as provided by the claimed invention set forth in claims 1 and 7.

Thus, a combination of the teachings of the Hendrickson paper, the Everett paper and the Andrea paper would not render unpatentable the claimed invention.

Should the Examiner disagree therewith, it is respectfully requested that the Examiner specify where in the cited document there is a basis for such disagreement.

Regarding claims 3, 5 and 6, applicants respectfully point out that claims 3, 5 and 6 depend on and include all the limitations of claim 1. Thus, claims 3, 5 and 6 are patentable at least for the reasons set forth above with respect to claim 1.

Regarding claims 9, 11 and 12, applicants respectfully point out that claims 9, 11 and 12 depend on and include all the limitations of claim 7. Thus, claims 9, 11 and 12 are patentable at least for the reasons set forth above with respect to claim 7.

Accordingly, applicants respectfully request reconsideration and withdrawal of the rejection of claims 1, 3, 5-7, 9, 11 and 12 under 35 U.S.C. §103(a).

Rejection Under 35 U.S.C. § 103(a)

In Section 6 of the June 18, 2001 Office Action, claims 4 and 10 were rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the Bachar paper, the Hendrickson paper, the Everett paper and the Andrea paper, as applied to claims 1 and 7, and further in view of Lima et al., "MAD analysis of FHIT, a putative human tumor suppressor from the HIT protein family," Structure 5, 763-774, 1997 (hereinafter "Lima paper").

The June 18, 2001 Office Action stated that the focus of the

Bachar paper is in proteomics wherein protein classification is based on the x-ray crystallographic data, and not in the instrumental improvements of x-ray crystallography. The Office Action alleged that one of ordinary skill in the art would recognize that improvement analytical methods as taught by the Lima paper for obtaining data with higher precision and/or resolution would be beneficial to the bioinformatics tool of the Bachar paper. The Office Action stated that the Lima paper teaches using an undulator beamline x-ray source, with MAD because of the high output levels, with narrow, tunable, harmonic peaks.

The June 18, 2001 Office Action alleged that it would have been obvious to one of ordinary skill in the art at the time the claimed invention was made to utilize the undulator beamline x-ray source, in place of the synchrotron device as taught by the Hendrickson paper, because the Lima paper demonstrates a high output x-ray source, with narrow, tunable, harmonic peaks.

Applicants maintain that the Bachar paper, the Hendrickson paper, the Everett paper, the Andrea paper and the Lima paper do not render unpatentable the claimed invention set forth in claims 4 and 10. The claimed invention is patentable over the Bachar paper, the Hendrickson paper, the Everett paper, the Andrea paper and the Lima paper for at least the following reasons.

As stated above, the Bachar paper teaches away from use of sequence information to compare/cluster protein structures and relates only to that subset of all proteins for which three-diemnsional structures are already known. Therefore, one skilled in the art would not have been motivated to combine the teachings of the Bachar paper with teachings of use of sequence data, for example, in the Everett paper.

The additional references cited in the Office Action, the Hendrickson paper, the Everett paper, the Andrea paper and the Lima paper, do not supply the motivation and information that the Bachar paper lacks, and therefore do not, alone or in combination, render obvious claim 4 and claim 10 which depend from claims 1 and 7, respectively.

As stated above, the Hendrickson paper, the Everett paper and the Andrea paper fail to describe or suggest the benefits of, for example, (a) systematically organizing in a comprehensive database of sequence information for a plurality of proteins and structural information and functional information for selected proteins, (b) using the sequence information, structural information and functional information in the database to cluster the plurality of proteins into families of proteins, such that family members have homologous sequences, and (c) for each family, selecting a plurality of family members as target proteins for structure determination, as provided by the claimed invention set forth in claims 1 and 7 from which claims 4 and 10 depend, respectively.

The Lima paper relates to a process which uses MAD (multiwavelength anomalous diffraction) analysis for determining the three-dimensional structure of a select protein [i.e. fragile histidine triad (FHIT)]. The Lima paper was cited in the Office Action for its description of use of an undulator beamline x-ray source with MAD.

While the Lima paper relates to structure determination of a select protein, the Lima paper, like the Hendrickson paper, the Everett paper and the Andrea paper, does not describe or suggest, however, (a) systematically organizing in a comprehensive database of sequence information for a plurality of proteins and

structural information and functional information for selected proteins, (b) using the sequence information, structural information and functional information in the database to cluster the plurality of proteins into families of proteins, such that family members have homologous sequences, and (c) for each family, selecting a plurality of family members as target proteins for structure determination, as provided by the claimed invention set forth in claims 1 and 7 from which claims 4 and 10 depend, respectively.

Therefore, the Bachar paper, the Hendrickson paper, the Everett paper, the Andrea paper and the Lima paper do not render obvious the claimed invention.

Accordingly, applicants respectfully request reconsideration and withdrawal of the rejection of claims 4 and 10 under 35 U.S.C. §103(a).

Rejection Under 35 U.S.C. § 103(a)

In Section 7 of the June 18, 2001 Office Action, claims 2 and 8 were rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over the Bachar paper, the Hendrickson paper, the Everett paper, and the Andrea paper, as applied to claims 1 and 7, and further in view of U.S. Patent No. 5,525,198 to Craig et al. (hereinafter "the Craig patent").

The June 18, 2001 Office Action stated that the focus of the Bachar paper is in proteomics wherein protein classification is based on the x-ray crystallographic data, and not in the instrumental improvements of x-ray crystallography. The Office Action alleged that one of ordinary skill in the art would recognize that improvement analytical methods as taught by the Craig patent for obtaining data with higher precision and/or

resolution would be beneficial to the bioinformatics tool of the Bachar paper. The Office Action stated that the Craig patent teaches the cryogenic freezing of target protein crystals as a means of increasing the crystal's stability during exposure to x-ray sources.

The June 18, 2001 Office Action alleged that it would have been obvious to one of ordinary skill in the art at the time the invention was made to utilize a means for the cryogenic cooling of the target protein crystal, with the system/process of the Bachar paper, in view of the Hendrickson paper, as the Craig patent teaches that cryogenic cooling preserves crystals during x-ray sampling. The Office Action also alleged that the invention as a whole was prima facie obvious at the time the invention was made.

Applicants maintain that the Bachar paper, the Hendrickson paper, the Everett paper, the Andrea paper and the Craig patent do not render unpatentable the claimed invention set forth in claims 2 and 8. The claimed invention is patentable over the Bachar paper, the Hendrickson paper, the Everett paper, the Andrea paper and the Craig patent for at least the following reasons.

As stated above, the Bachar paper teaches away from use of sequence information to compare/cluster protein structures and relates only to that subset of all proteins for which three-dimensional structures are already known. Therefore, one skilled in the art would not have been motivated to combine the teachings of the Bachar paper with teachings of use of sequence data, for example, in the Everett paper.

The additional references cited in the Office Action, the Hendrickson paper, the Everett paper, the Andrea paper and the

Craig patent, do not supply the motivation and the information that the Bachar paper lacks, and therefore do not, alone or in combination, render obvious claim 2 and claim 8 which depend from claim 1 and claim 7, respectively.

As stated above, the Hendrickson paper, the Everett paper and the Andrea paper fail to describe or suggest the benefits of, for example, (a) systematically organizing in a comprehensive database of sequence information for a plurality of proteins and structural information and functional information for selected proteins, (b) using the sequence information, structural information and functional information in the database to cluster the plurality of proteins into families of proteins, such that family members have homologous sequences, and (c) for each family, selecting a plurality of family members as target proteins for structure determination, as provided by the claimed invention set forth in claims 1 and 7 from which claims 2 and 8 depend, respectively.

The Craig patent relates to determination of the 3-D structure of a molecule by forming an electrorheological crystalline mass of the molecule, obtaining an X-ray diffraction pattern of the electrorheological crystalline mass, and calculating the 3-D structure of the molecule using the X-ray diffraction pattern. The Craig patent was cited in the Office Action for its description of cryogenic freezing of target protein crystals as a means of increasing the crystal's stability during exposure to x-ray sources.

The Craig patent, like the Hendrickson paper, the Everett paper and the Andrea paper, does not describe or suggest, however, (a) systematically organizing in a comprehensive database of sequence information for a plurality of proteins and structural

information and functional information for selected proteins, (b) using the sequence information, structural information and functional information in the database to cluster the plurality of proteins into families of proteins, such that family members have homologous sequences, and (c) for each family, selecting a plurality of family members as target proteins for structure determination, as provided by the claimed invention set forth in claims 1 and 7 from which claims 2 and 8 depend, respectively.

Therefore, the Bachar paper, the Hendrickson paper, the Everett paper, the Andrea paper and the Craig patent do not render obvious the claimed invention.

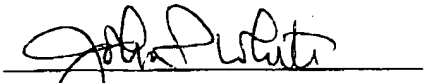
Accordingly, applicants respectfully request reconsideration and withdrawal of the rejection of claims 2 and 8 under 35 U.S.C. §103(a).

In view of the amendments to the claims and remarks hereinabove, applicants maintain that claims 1 through 12 are now in condition for allowance. Accordingly, applicants earnestly solicit the allowance of claims 1 through 12.

If an interview would be of assistance in advancing prosecution of the subject application, applicants' undersigned attorney invites the Examiner to telephone him at the telephone number provided below.

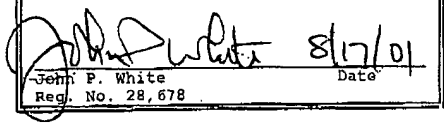
No fee is deemed necessary in connection with the filing of this Amendment. However, if any fee is required, authorization is hereby given to charge the amount of any such fee to Deposit Account No. 03-3125.

Respectfully submitted,



John P. White
Registration No. 28,678
Attorney for Applicants
Cooper & Dunham LLP
1185 Avenue of the Americas
New York, New York 10036
(212) 278-0400

I hereby certify that this correspondence is being transmitted this date by facsimile to the Patent And Trademark Office and deposited this date with the U.S. Postal Service with sufficient postage as first class mail in an envelope addressed to: Assistant Commissioner for Patents, Box AF, Washington, D.C. 20231.



8/17/01
John P. White Date
Reg. No. 28,678

Wayne A. Hendrickson and Barry Honig
Serial No.: 09/235,986
Filed: January 22, 1999

Exhibit A

(AMENDMENT UNDER 37 C.F.R. § 1.116 IN
RESPONSE TO JUNE 18, 2001 FINAL OFFICE ACTION)

1. (Four Times Amended) A system for determining experimentally a plurality of three-dimensional atomic structures, each of which is associated with a corresponding protein, comprising:

a database of sequence information for a plurality of proteins, and structural information and functional information for selected proteins;

at least one bioinformatics tool adapted to use the sequence information, structural information and functional information stored in the database to cluster the plurality of proteins into a plurality of families, in which, for each family, members of the family have [corresponding] homologous sequences;

protein synthesis means for synthesizing for each family determined by the at least one bioinformatics tool a plurality of target proteins which are members of the family, using information stored in the database corresponding to the target proteins, the protein synthesis means having screening means for screening products of the synthesis to choose selected synthesized products for processing;

protein processing means for preparing, purifying and characterizing each of the selected synthesized products;

crystallization means for crystallizing the processed synthesized product against a plurality of crystallization screens to produce a plurality of specimen crystals of the target protein, and testing the plurality of specimen crystals for predetermined diffraction characteristics to determine suitable specimen crystals;

X-ray crystallography means for performing high-throughput crystallography on the specimen crystals of each target protein determined by the crystallization means to be suitable, the X-ray crystallography means having diffraction measuring means for measuring for diffraction data the suitable specimen crystals of

the target protein, analyzing means for analyzing the diffraction data, means for building an atomic model of the target protein according to an analysis of the diffraction data by the analyzing means, and means for refining the model of the target protein against the diffraction data and storing the refined model in the database;

structure extraction means having means for analyzing the refined model of the target protein using sequence information corresponding to other family members which is stored in the database and structural information corresponding to other proteins which is stored in the database, means for analyzing the refined model for functional motifs and for surface characteristics to define active sites and macromolecular contact sites, and means for defining at least one class of compounds predicted to have binding potency using the active sites information corresponding to the target protein; and

a homology model building tool adapted to use the refined model of the target protein retrieved from the database to develop a homology model of one or more predicted protein structures,

wherein the database is updated using the at least one bioinformatics tool and the developed homology model.

7. (Four Times Amended) A process for determining experimentally a plurality of three-dimensional atomic structures, each of which is associated with a corresponding protein, comprising the steps of:

(a) systematically organizing sequence information for a plurality of proteins, and structural information and functional information for selected proteins into a database;

(b) clustering the plurality of proteins into a plurality of families, in which, for each family, members of the family

have [corresponding] homologous sequences, using at least one bioinformatics tool and the sequence information, structural information and functional information stored in the database;

(c) synthesizing for each family determined in step (b) a plurality of target proteins which are members of the family, using information stored in the database corresponding to the plurality of target proteins, and screening products of the synthesis to choose selected synthesized products for processing;

(d) preparing, purifying and characterizing each synthesized product that is chosen in step (c);

(e) crystallizing the processed synthesized product prepared, purified and characterized in step (d) against a plurality of crystallization screens to produce a plurality of specimen crystals of the target protein;

(f) testing the plurality of specimen crystals grown in step (e) for predetermined diffraction characteristics to determine suitable specimen crystals of the target protein;

(g) performing high-throughput crystallography, including measuring for diffraction data the specimen crystals determined in step (f) to be suitable, building an atomic model of the target protein according to an analysis of the diffraction data, refining the model of the target protein against the diffraction data, and storing the refined model in the database;

(h) analyzing the refined model, stored in the database in step (g), of the target protein using sequence information corresponding to other family members which is stored in the database and structural information corresponding to other proteins which is stored in the database, analyzing the refined model of the target protein for functional motifs and for surface characteristics to define active sites and macromolecular contact sites, and defining at least one class of compounds predicted to have binding potency using the active sites information

corresponding to the target protein;

(i) developing a homology model of one or more predicted protein structures using computational tools for homology model building and the refined model of the target protein retrieved from the database, and updating the database by using the at least one bioinformatics tool and the developed homology model; and

(j) performing steps (f) through (i) for each of the other target proteins.

Wayne A. Hendrickson and Barry Honig
Serial No.: 09/235,986
Filed: January 22, 1999

Exhibit B

(AMENDMENT UNDER 37 C.F.R. § 1.116 IN
RESPONSE TO JUNE 18, 2001 FINAL OFFICE ACTION)

1. (Four Times Amended) A system for determining experimentally a plurality of three-dimensional atomic structures, each of which is associated with a corresponding protein, comprising:

a database of sequence information for a plurality of proteins, and structural information and functional information for selected proteins;

at least one bioinformatics tool adapted to use the sequence information, structural information and functional information stored in the database to cluster the plurality of proteins into a plurality of families, in which, for each family, members of the family have homologous sequences;

protein synthesis means for synthesizing for each family determined by the at least one bioinformatics tool a plurality of target proteins which are members of the family, using information stored in the database corresponding to the target proteins, the protein synthesis means having screening means for screening products of the synthesis to choose selected synthesized products for processing;

protein processing means for preparing, purifying and characterizing each of the selected synthesized products;

crystallization means for crystallizing the processed synthesized product against a plurality of crystallization screens to produce a plurality of specimen crystals of the target protein, and testing the plurality of specimen crystals for predetermined diffraction characteristics to determine suitable specimen crystals;

X-ray crystallography means for performing high-throughput crystallography on the specimen crystals of each target protein determined by the crystallization means to be suitable, the X-ray crystallography means having diffraction measuring means for measuring for diffraction data the suitable specimen crystals of

the target protein, analyzing means for analyzing the diffraction data, means for building an atomic model of the target protein according to an analysis of the diffraction data by the analyzing means, and means for refining the model of the target protein against the diffraction data and storing the refined model in the database;

structure extraction means having means for analyzing the refined model of the target protein using sequence information corresponding to other family members which is stored in the database and structural information corresponding to other proteins which is stored in the database, means for analyzing the refined model for functional motifs and for surface characteristics to define active sites and macromolecular contact sites, and means for defining at least one class of compounds predicted to have binding potency using the active sites information corresponding to the target protein; and

a homology model building tool adapted to use the refined model of the target protein retrieved from the database to develop a homology model of one or more predicted protein structures,

wherein the database is updated using the at least one bioinformatics tool and the developed homology model.

7. (Four Times Amended) A process for determining experimentally a plurality of three-dimensional atomic structures, each of which is associated with a corresponding protein, comprising the steps of:

(a) systematically organizing sequence information for a plurality of proteins, and structural information and functional information for selected proteins into a database;

(b) clustering the plurality of proteins into a plurality of families, in which, for each family, members of the family

have homologous sequences, using at least one bioinformatics tool and the sequence information, structural information and functional information stored in the database;

(c) synthesizing for each family determined in step (b) a plurality of target proteins which are members of the family, using information stored in the database corresponding to the plurality of target proteins, and screening products of the synthesis to choose selected synthesized products for processing;

(d) preparing, purifying and characterizing each synthesized product that is chosen in step (c);

(e) crystallizing the processed synthesized product prepared, purified and characterized in step (d) against a plurality of crystallization screens to produce a plurality of specimen crystals of the target protein;

(f) testing the plurality of specimen crystals grown in step (e) for predetermined diffraction characteristics to determine suitable specimen crystals of the target protein;

(g) performing high-throughput crystallography, including measuring for diffraction data the specimen crystals determined in step (f) to be suitable, building an atomic model of the target protein according to an analysis of the diffraction data, refining the model of the target protein against the diffraction data, and storing the refined model in the database;

(h) analyzing the refined model, stored in the database in step (g), of the target protein using sequence information corresponding to other family members which is stored in the database and structural information corresponding to other proteins which is stored in the database, analyzing the refined model of the target protein for functional motifs and for surface characteristics to define active sites and macromolecular contact sites, and defining at least one class of compounds predicted to have binding potency using the active sites information

corresponding to the target protein;

(i) developing a homology model of one or more predicted protein structures using computational tools for homology model building and the refined model of the target protein retrieved from the database, and updating the database by using the at least one bioinformatics tool and the developed homology model; and

(j) performing steps (f) through (i) for each of the other target proteins.